Discrete Mathematics 50 (1984) 285–293
North-Holland


# A GENERALIZED CLOSURE AND COMPLEMENT PHENOMENON

D. PELEG

*Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel*

Received 3 May 1983

The number of different sets that can be generated from a given set by applications of complement and closure operators is finite and small (e.g., 14). This fact, stated originally in [4] for topological closures, and later in [2] for transitive closures of binary relations, is generalized to other closure operators, and several examples are given.

## 1. Introduction

Let $X$ be a set, and $f_1, \ldots, f_k : 2^X \to 2^X$ some operators. Given a set $A \subseteq X$, it is possible to apply various elements from among $f_1, \ldots, f_k$ to $A$, yielding different sets such as $A^{f_1 f_1 f_2}$, $A^{f_3 f_2 f_1 f_3}$ etc. (where $A^{fg}$ is a shorthand for $(A^f)^g$). Denote by $G(f_1, \ldots, f_k, A)$ the closure of $A$ under $f_1, \ldots, f_k$ (i.e. the collection of sets which can be generated from $A$ in such ways). Next, denote $g(f_1, \ldots, f_k) = \max_{A \subseteq X}\{|G(f_1, \ldots, f_k, A)|\}$. This paper deals with the possible values of $g$ in the limited case in which $k = 2$, and the two operators are complementation with respect to $X$, denoted by $A^-$, and some operator $\circ$ satisfying certain natural closure properties.

All closure operators considered here are idempotent, (i.e. for all $A \subseteq X$, $A^{\circ\circ} = A^\circ$), and since $A^{--} = A$, it is clear that the value of $g$ is determined by the behavior of the series obtained by alternating applications of $\circ$ and $-$ (i.e., $A$, $A^\circ$, $A^{\circ -}$, $A^{\circ - \circ}$ etc.). The series can either become cyclic at some stage, or alternatively keep on generating new sets. However, the phenomenon investigated in this paper is that for *closure operators* satisfying some natural properties, the series always becomes cyclic at an early stage, and the number of different sets is therefore very small (e.g., 6, 10 or 14).

This phenomenon was first observed by Kuratowski [4], with regard to topological closures. A *topological closure* is an operator $\circ : 2^X \to 2^X$ which satisfies the following 'Kuratowski axioms': For $A, B \subseteq X$,

(K1) $A \subseteq A^\circ$,
(K2) $A^{\circ\circ} = A^\circ$,
(K3) $A^\circ \cup B^\circ = (A \cup B)^\circ$,
(K4) $\emptyset^\circ = \emptyset$.

Kuratowski also observed the correspondence between topological closures and topologies; Every topological closure $\circ$ determines a unique topology defined by $\mathcal{T} = \{A^{\circ-} \mid A \subseteq X\}$, and, conversely, every topology $\mathcal{T}$ determines a unique closure operator defined by $A^{\circ} = \bigcap \{B \mid A \subset B \subseteq X, B^{-} \in \mathcal{T}\}$.

Kuratowski proved that for any topological closure $\circ$, and any set $A$, $A^{\circ-\circ-\circ-\circ} = A^{\circ-\circ}$, and therefore at most 14 different sets can be generated from a given set. He has also exhibited, for the usual topology on the reals, a specific set which achieves this upper bound.

It was not until some decades later, that the same phenomenon was observed in a different context – for *transitive closures* of binary relations. Let $X = D \times D$ for some domain $D$. A relation $R \subseteq X$ is *transitive* if it obeys the following inclusion rule: If $\langle a, b\rangle, \langle b, c\rangle \in R$ then $\langle a, c\rangle \in R$. The transitive closure $R^{+}$ of a relation $R$ is defined to be the smallest transitive relation containing $R$. Graham, Knuth and Motzkin [2] show that for any binary relation $R$, $R^{+-+-+} = R^{+-+-}$ (or as they phrase it: "the complement of the transitive closure of the complement of a transitive relation is transitive"), and hence $g(-,+) \leqslant 10$. They also analyze the possible values of $|G(-,+,R)|$ according to the structure of $R$, and exhibit a relation which actually produces 10 distinct relations, so that $g(-,+) = 10$.

Though Graham et al. point out the analogy with Kuratowski's result, they do not pursue the connection any further. Note, however, that even the bound of 14 cannot be derived from Kuratowski's result, since transitive closure is not a topological closure. In particular, it does not obey Axiom K3.

These results have recently gained some momentum in connection with the theory of relational data bases (which are, essencially, the computer science jargon for finite structures). Chandra and Harel [1] analyze the structural complexity of queries (i.e., functions) on relational data bases, by defining certain hierarchies of complexity levels for queries. It is still an open question whether some of the levels in these hierarchies are distinct from one another and an earlier version of [1] proposed using repetitive applications of complementation and j-fold transitive closure to show the strictness of a certain infinite sub-hierarchy, (a j-fold transitive closure is a generalization of the usual transitive closure to relations of rank j). Of course, if j-fold transitive closure turns out to exhibit the same behavior as the usual (binary) transitive closure (as indeed will be shown here), then this possibility is ruled out, since there will be only finitely many distinct sets available.

In this paper we extend the *closure and complement* result to several closure operators which we categorize as *semi-topological* closures. These operators do not satisfy all four of Kuratowski's axioms, yet their properties suffice to maintain this phenomenon. Following the definitions in Section 2, we analyze in Section 3 the typical behavior of such operators, give the bound of 14 distinct sets and define a subclass of compact operators for which the bound is 10. The transitive closure of binary relations belongs to this subclass. Some necessary and sufficient conditions for the latter subclass are given. Section 4 contains several examples. Finally, in Section 5 we suggest directions for further research.

## 2. Closures and semi-topologies

We start by defining a special class of operators $\circ: 2^X \to 2^X$ called *semi-topological closures*. These are operators satisfying the following three axioms: For $A, B \subseteq X$,

(S1) $A \subseteq A^\circ$,
(S2) $A^{\circ\circ} = A^\circ$, (idempotence)
(S3) $A^\circ \cup B^\circ \subseteq (A \cup B)^\circ$ (monotonicity)

(S3 can be expressed equivalently as $A \subseteq B \Rightarrow A^\circ \subseteq B^\circ$).
Next define a *semi-topology* $\mathcal{T}$ on $X$, by:

(T1) $\mathcal{T} \subseteq 2^X$,
(T2) $\emptyset \in \mathcal{T}$,
(T3) $\mathcal{T}$ is closed under (finite or infinite) unions.

As with topologies, define: A set $A$ is *open* if $A \in \mathcal{T}$. $A$ is closed if $A^-$ is open. Denote the collection $\{A^- \mid A \in \mathcal{T}\}$ of closed sets by $C(\mathcal{T})$.

The correspondence between semi-topologies and semi-topological closures is analogous to that between topologies and topological closures:

**Lemma 2.1.** *There is a 1–1 onto correspondence between semi-topologies and semi-topological closures. For each semi-topology $\mathcal{T}$ there is a unique semi-toplogical closure $\circ$ so that*

(1) $\mathcal{T} = \{A^{\circ -} \mid A \subseteq X\}$,
(2) $A^\circ = \bigcap \{B \mid A \subseteq B \subseteq X, B \in C(\mathcal{T})\}$.

Thus one can define $^\circ$ either by its operation or by characterizing the open (or closed) sets in its corresponding semi-topology.

A common way of defining semi-topological closure operators is by means of inclusion rules, which are general schemas of the form:

$$F(A): B \subseteq A \Rightarrow C \subseteq A.$$

Such schemas can be utilized in two ways:

(1) Characterization of closed sets: Define $A$ as closed iff all instances of $F(A)$ are true. Proceed to define a semi-topology and a semi-topological closure as in Lemma 2.1.

(2) Description of the 'closing' process: Obtain $A^\circ$ by applying $F$ (read as: "If $B \subseteq A$ then include also $C$ in $A$") until nothing new can be added.

It is straightforward to verify that both approaches indeed produce a semi-topological closure on $X$.

In some cases, the inclusion rule has the simple structure:

$$F(A): a, b \in A \Rightarrow c \in A.$$

In such cases one can 'realize' $F$, using a partial binary operator $\cdot : X \times X \to X$.

This operator can be extended to sets on $X$, i.e. $A \cdot B = \{a \cdot b \mid a \in A \wedge b \in B\}$. Finally, let $f(A) = A \cup A \cdot A$, and let $A^\circ = \lim(f^k(A))$. (Note that the series $A, f(A), f(f(A)), \ldots$ is monotonically increasing and bounded by $X$). When $\cdot$ is associative, one can naturally define $A^i = A \cdot \ldots \cdot A$, $i$ times, and clearly $f^k(A) = \bigcup_{i=1}^{2^k} A^i$. Therefore it is possible to define $A^\circ$ by $\bigcup_{i=1}^{\infty} A^i$. We call such an operator $\circ$ a *power series* operator.

**Example.** Transitive closure of binary relations (subsets of $X = D \times D$), is defined by using the inclusion rule: $\langle a, b\rangle, \langle b, c\rangle \in R \Rightarrow \langle a, c\rangle \in R$. This rule was used earlier for characterizing the closed (transitive) relations. On the other hand, one can define the transitive closure directly, as the power series operator based on the associated realizing operator defined by $\langle a, b\rangle \cdot \langle b, c\rangle = \langle a, c\rangle$.

As a final remark, it should be mentioned that analogous to semi-topological closures, one can define *semi-topological interiors*, by replacing Axiom S1 with

(S1′) $A \supseteq A^\circ$.

The behaviour of interiors (with respect to the discussed phenomenon) mostly parallels that of closures, so we shall not bother to describe it separately.

## 3. 'Big bear' and 'little bear' schemata

This section contains the closure and complement results for semi-topological closures.

**Lemma 3.1.** *Let $\circ$ be a semi-topological closure, and let $A \subseteq X$. Then $A^{\circ-\circ-\circ-\circ} = A^{\circ-\circ}$.*

**Proof.** Substituting $A^{\circ-}$ for $A$ in Axiom S1, yields $A^{\circ-} \subseteq A^{\circ-\circ}$. Now apply $-$ and then $\circ$ to both sides and obtain $(*)$ $A^{\circ-\circ-\circ} \subseteq A^\circ$. Now:

(1) Substitute $A^{\circ-}$ for $A$ in $(*)$ to yield $A^{\circ-\circ} \supseteq A^{\circ-\circ-\circ-\circ}$,

(2) Apply $-\circ$ to both sides of $(*)$ to yield $A^{\circ-\circ} \subseteq A^{\circ-\circ-\circ-\circ}$.

Combining (1) and (2) yields the desired property. □

It is now possible to sketch the process of applying $\circ$ and$-$ alternately to a given set $A$, as in Fig. 1. Note that the *big bear* schema covers all possibilities, since

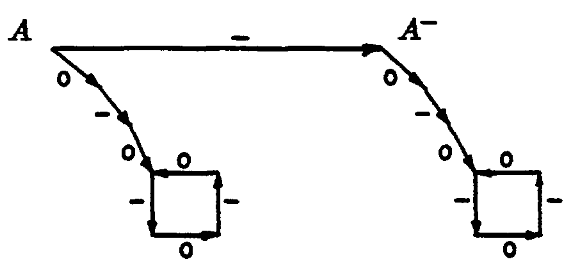


Fig. 1. Big bear schema.

$A^{--}=A$ and $A^{\circ\circ}=A^{\circ}$. Therefore

**Corollary 3.2.** *For a semi-topological closure* $\circ$, $g(-,\circ)\leqslant 14$.

Let us examine the schema more closely. Let $S=A^{\circ-\circ-}$, $Q=S^{\circ-}$, and $\Delta=S^{\circ}-S$. Since $Q^{\circ-}=S$, we have $Q^{\circ}-Q=S^{-}-S^{\circ-}=S^{\circ}-S=\Delta$. Also $S\cap\Delta=Q\cap\Delta=S\cap Q=\emptyset$ and $S\cup Q\cup\Delta=X$. All this is summarized in Fig. 2.

Clearly, if $\Delta$ is $\emptyset$ for some $A$, then $A^{\circ-\circ-\circ}=A^{\circ-\circ-}$, and when this is the case for every $A\subseteq X$, the schema becomes that of Fig. 3.

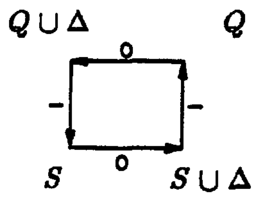


Fig. 2.

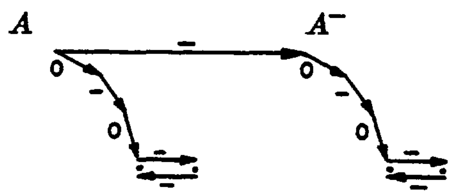


Fig. 3. Little bear schema.

**Definition 3.3.** A semi-topological closure $\circ$ is *compact* if

$$\forall A\subseteq X(A^{\circ-\circ-\circ}=A^{\circ-\circ-}).$$

**Corollary 3.4.** *For a compact operator* $\circ$, $g(-,\circ)\leqslant 10$.

We now discuss some necessary and sufficient conditions for compactness.

**Lemma 3.5.** *If* $\circ$ *is a compact closure operator then* $\emptyset^{\circ}=\emptyset$.

**Proof.** Assume $\emptyset^{\circ}\neq\emptyset$, or, $\exists a\ (a\in\emptyset^{\circ})$. Then $\forall A\subseteq X$, $a\in\emptyset^{\circ}\subseteq\emptyset^{\circ}\cup A^{\circ}\subseteq(\emptyset\cup A)^{\circ}=A^{\circ}$. Therefore $\forall A\subseteq X$, $a\in A^{\circ-\circ}$, $a\notin A^{\circ-\circ-}$ and $a\in A^{\circ-\circ-\circ}$. This implies $A^{\circ-\circ-}\neq A^{\circ-\circ-\circ}$, so $\circ$ is not compact. $\square$

The following definition gives a topological version for the compactness property.

**Definition 3.6.** A semi-topological closure $\circ$ is *openness preserving* if in the semi-topology defined by $\circ$, $\forall A$ ($A$ is open $\Rightarrow A^{\circ}$ is open), or, $\forall A$ $(A^{-\circ}=A^{-}\Rightarrow A^{\circ-\circ}=A^{\circ-})$.

**Lemma 3.7.** *Let* $\circ$ *be a semi-topological closure. Then* $\circ$ *is compact iff it is openness preserving.*

**Proof.** ($\Rightarrow$): Assume $\circ$ is compact, but is not openness preserving. That is, for some $A\subseteq X$, (*) $A^{-}=A^{-\circ}$, yet $A^{\circ-}\subset A^{\circ-\circ}$, or, (**) $A^{\circ-\circ-}\subset A^{\circ}$. Applying $-\circ$ to

Axiom S1 and using (*) gives $A^{\circ-\circ} \subseteq A^{-\circ} = A^-$, or $A \subseteq A^{\circ-\circ-}$. Together with (**) we have $A \subseteq A^{\circ-\circ-} \subset A^\circ$, which implies (by applying $\circ$ to all terms) $A^{\circ-\circ-\circ} = A^\circ$. Therefore $A^{\circ-\circ-\circ} \neq A^{\circ-\circ-}$, contradicting the assumption.

($\Leftarrow$): Assume $\circ$ is openness preserving. For any $A$, $(A^{\circ-})^{-\circ} = A^\circ = (A^{\circ-})^-$, so by the assumption $(A^{\circ-})^{\circ-\circ} = (A^{\circ-})^{\circ-}$. $\square$

For the subclass of power series operators defined earlier, the same conditions can be stated in more suitable forms. It is known that a set $A$ closed under a power series operator (i.e. $A = A^\circ$) iff $A^2 \subseteq A$. Therefore, the openness preserving condition for such operators takes the form

$$\forall A \subseteq X((A^-)^2 \subseteq A^- \Rightarrow (A^{\circ-})^2 \subseteq A^{\circ-}).$$

**Definition 3.8.** A power series operator $\circ$ has the *decomposition property* if

$$\forall u, v, w_1, \ldots, w_k \in X(u \cdot v = w_1 \cdot \ldots \cdot w_k \Rightarrow \exists i, \exists x_1, x_2 \in X,$$
$$(1 \leqslant i \leqslant k \wedge u = w_1 \cdot \ldots \cdot w_{i-1} \cdot x_1 \wedge v = x_2 \cdot w_{i+1} \cdot \ldots \cdot w_k)).$$

Now the following holds:

**Lemma 3.9.** *Let $\circ$ be a power series operator with the decomposition property. Then $\circ$ is compact.*

**Proof.** Assume $(A^{\circ-})^2 \subseteq A^{\circ-}$ does not hold. That is, $\exists u, v$ $(u, v \in A^{\circ-} \wedge u \cdot v \notin A^{\circ-})$ for some $A \subseteq X$. Therefore $u \cdot v \in A^\circ$, which implies $\exists w_1, \ldots, w_k$ $(u \cdot v = w_1 \cdot \ldots \cdot w_k \wedge \forall 1 \leqslant i \leqslant k$ $(w_i \in A))$. By the decomposition property of $\circ$, $\exists i, \exists x_1, x_2 \in X$ $(1 \leqslant i \leqslant k \wedge u = w_1 \cdot \ldots \cdot w_{i-1} \cdot x_1 \wedge v = x_2 \cdot w_{i+1} \cdot \ldots \cdot w_k))$. Clearly $x_1, x_2 \notin A$ (otherwise $u$ or $v$ are in $A^\circ$), yet $x_1 \cdot x_2 = w_i \in A$. Therefore $x_1 \cdot x_2 \in (A^-)^2$ and $x_1 \cdot x_2 \notin A^-$, hence $(A^-)^2 \subseteq A^-$ does not hold either. $\square$

## 4. Examples

In this section we examine some specific semi-topological operators, and indicate the implications of our results.

To begin with, all topological closure operators are in particular semi-topological closures, so Corollary 3.2 extends Kuratowski's 'closure and complement' theorem. The same applies also to the transitive closure for binary relations, as it satisfies Axioms S1 to S3. This suffices to get $g(-,+) \leqslant 14$. Furthermore, as transitive closure is a power series operator, we get the improved bound of [2] from Lemma 3.9, since transitive closure possesses the decomposition property. Other examples are listed below.

## 4.1. *Kleene closures*

Let $\Sigma^*$ be the set of all finite strings over an alphabet $\Sigma$, and define concatenation as usual. Then *Kleene* $\oplus$ is the power series operator based on concatenation. Next, let $L^0=\{\lambda\}$ for every $L\subseteq\Sigma^*$, where $\lambda$ is the empty string, and $L^*=L^{\oplus}\cup L^0$ $(=\bigcup_{i=0}^{\infty}L^i)$. This operator is known as the *Kleene star.* Both Kleene operators are semi-topological closures, hence the upper bound $g(-,*)$, $g(-,\oplus)\leqslant 14$. By Lemma 3.6, $*$ is not compact (since $\emptyset^*=\{\lambda\}$). Indeed, the upper bound can be shown tight. For example, for $L=\{a, aab, bbb\}$ we get $|G(-,*,L)|=14$.

On the other hand, Kleene $\oplus$ is a power series operator, and it has the decomposition property, so by Lemma 3.9 $g(-,\oplus)\leqslant 10$. The same language $L$ established 10 as the exact bound.

## 4.2. *Constant union and intersection*

The following examples illustrate the fact that there are non-compact operators that generate less than 14 sets. Let us define $\sqcup$ by $A^{\sqcup}=A\cup C$ for a specific $C$, $\emptyset\subset C\subset X$, and similarly define $\sqcap$ by $A^{\sqcap}=A\cap C$. It is easy to verify that $\sqcup$ is a semi-topological closure and $\sqcap$ is a semi-topological interior. By Lemma 3.5 and its analog for interiors, both $\sqcup$ and $\sqcap$ are not compact, yet it is clear that $g(-,\sqcup)=g(-,\sqcap)=6$ (and even 4 when $C$ is a singleton).

One special case of a $\sqcup$ operator is the *reflexive closure* $I$ defined over $X=D\times D$ by $R^I=A\cup\Delta$ where $\Delta=\{(d,d)\mid d\in D\}$. By the above remark, $g(-,I)=6$. In [2] it is shown that $g(-,+,I)=42$.

## 4.3. *Transitive closures of non-binary relations*

Transitive closure of binary relations was defined in the previous sections by means of inclusion rules, which were used to construct a 'power series' definition. Generalization to non-binary relations gets obscured for several reasons. To begin with, there is no clear concept of transitive closure for the higher ranks, and many 'natural' definitions are possible. Secondly, in many such attempts of generalization, it becomes cumbersome to define a power series operator, as the desired inclusion rule (of the form: If $B\subseteq A$ then also $C\subseteq A$), does not necessarily satisfy $|B|=2$ and $|C|=1$. Therefore, the associated 'realizing' operator $\cdot$ has to be of the form $\cdot:2^X\to 2^X$, and such a characterization is further complicated when this operator is not associative.

It should be clear, however, that such closures – even though they cannot be described 'conveniently' as power series operators – can be computed in a very similar manner to the computation of binary transitive closure.

Some possible inclusion rules (for ternary relations) are listed in Table 1. The definition given in [1] is TC1. The rules are expressed for ternary relations, for the sake of readability, and can be extended to higher ranks in the obvious way. Rule TC5 is due to Immerman [3]. It is defined only for relations of even rank, so the demonstration is for rank 4.

Table 1. Some possible definitions for a 3(4)-fold transitive relation

| Name | If $A$ contains | Then $A$ contains also |
|---|---|---|
| TC1 | *abc*, *bcd* | *abd*, *acd* |
| TC2 | *abc*, *cde* | *abe*, *ace*, *ade* |
| TC3 | *abc*, *cde* | *abd*, *abe*, *acd*, *ace*, *ade*, *bcd*, *bce*, *bde* |
| TC4 | *abc*, *bcd*, *cde* | *ace* |
| TC5 | *abcd*, *cdef* | *abef* |

It is easy to verify that the closure operators based on inclusion rules TC1 to TC5 are semi-topological closures. Hence, $g(-, \mathrm{TC}i) \leq 14$ for $i = 1, 2, 3, 4, 5$. This prevents such operations from being used directly to show strictness of the query hierarchies, as suggested in (an early version of) [1].

It is interesting to note that some seemingly natural attempts to define a transitive closure for non-binary relations lead to a very simple stabilization schema, in which always $R^{\circ-\circ} \in \{\emptyset, X\}$. This happens, for instance, for the closure operator based on inclusion rule TC3.

### 4.4. *Miscellaneous*

Graham et al. [2] consider also a situation in which $R \subseteq T$, where $T$ is some total order relation over $D$, and extend their result to cases where complementation is taken with respect to $T$ (instead of $D \times D$). It is obvious that any relation generated in this framework is contained in $T$ and that the transitive closure still enjoys the decomposition property within $T$, so the desired bound easily follows.

Another operator mentioned as worth considering (but not discussed explicitly) in [2] is the *difunctional closure* $D$ of a binary relation, defined by $R^D = (R \cdot R^T)^+ \cdot R$, where $R^T$ is the inverse of $R$. It is straightforward to verify that $D$ is a semi-topological closure, hence $g(-, D) \leq 14$ is immediate.

An operator which may be worth considering due to its possible uses in the theory of relational data bases is the projection on the first dimension, $\downarrow$ : For $R \subseteq D \times D$, $R^{\downarrow} = R \cdot (D \times D)$. Again, $\downarrow$ is a semi-topological closure.

## 5. Suggestions for further research

(1) Compact operators need a more useful characterization (equivalent to the openness preserving condition). In particular, the question of compactness is not settled for transitive closures of non-binary relations (by inclusion rule TC1, for instance).

(2) We have considered only values of $g(f_1, \ldots, f_k)$ for $k = 2$. Investigation of certain cases for more than two operators might be of interest.

(3) Studying the structure of the collection $G(f_1, \ldots, f_k, A)$ for particular sets

$A$ may lead to better understanding of the overall behaviour of $g(f_1, \ldots, f_k)$.

(4) It may be interesting to investigate what further properties of topologies can be generalized to semi-topologies. The concept of semi-topologies may in general be found useful.

## Acknowledgment

I wish to thank D. Harel for helpful discussions and suggestions during the preparation of this note.

## References

[1] A. Chandra and D. Harel, Structure and complexity of relational queries, J. Comput. System. Sci. 25 (1982) 99–128. (also: CS82-05, Weizmann Institute of Science, Dept. of Applied Mathematics, Rehovot, Israel).

[2] R.L. Graham, D.E. Knuth and T.S. Motzkin, Complements and transitive closures, Discrete Math. 2 (1972) 17–29.

[3] N. Immerman, Languages which capture complexity classes, 15th ACM SIGACT Symposium, April 1983.

[4] C. Kuratowski, Sur l'operation $\bar{A}$ de l'analysis situs, Fund. Math. 3 (1922) 182–199.